PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Which anthropometric measures best indicate type 2 diabetes among Russian, Somali and Kurdish origin migrants in Finland? A cross-sectional study.
AUTHORS	Skogberg, N.; Laatikainen, Tiina; Lundqvist, Annamari; Lilja, Eero; Härkänen, Tommi; Koponen, Päivikki

VERSION 1 – REVIEW

REVIEWER	Saula de Kreutzenberg University of Padova, Italy
REVIEW RETURNED	06-Sep-2017

GENERAL COMMENTS	The objective of the study by Skogberg et al. is to compare the performance of body mass index (BMI), waist-to-height ratio (WHtR), waist circumference (WC) and waist-to-hip ratio (WHR) in detecting type 2 diabetes among different migrant populations and Finnish. The Authors conclude that WC and WHtR were the best anthropometric measures for detecting type 2 diabetes among both Western and non-Western origin populations; however accuracy of detection was lower particularly among Somali and Kurds; diabetes risk assessment tools considering anthropometrics measures need to be validated among non-Western populations. The paper is well written, and the issue is interesting, since anthropometric measures represent an easy tool to consider for calculating the risk of diabetes. However, several limitations are linked to the study population. As acknowledged by the Authors, the number of participants is quite inadequate; the rate of subjects with >45 years is quite different among groups (61% in Russian, and 31% in Kurdish), and this may influence the results. The authors observed that abdominal obesity seems to have less importance in Somali and Kurds, in detecting diabetes, but the contributing role of abdominal obesity in diabetes development is not investigated in this study. We also lack information about the duration of diabetes, and the time of diagnosis, i.e. before or after the arrival in a Western country. This information would greatly improve the paper.

REVIEWER	Hongyan Xu Augusta University, United States
REVIEW RETURNED	28-Oct-2017

GENERAL COMMENTS	This is a study evaluating the performance of several anthropometric measures in detecting type 2 diabetes using a cross-section design. I have some concerns: 1. The Maamu survey used samples of 3000 subjects from a stratified random sample. The authors should clarified the

	stratification procedure. 2. In the statistical analysis section, the authors stated “all analyses accounted for the stratified sampling and finite population correction”. The methods used for achieving this should be specified rather than just specifying the software names. 3. On page 9, last line, the authors mentioned linear logistic regression was used to calculate age-adjusted mean values and 95% CI for continuous variables. This does not make sense. 4. On page 10, first two lines, the authors mentions regression analyses by study group stratified by sex. This is very confusing and needs to be clarified. In the result section, I did not see any results of the AUC by sex. 5. The last sentence of the method section, “With exception of inverse probability weights, the sampling design was not accounted for in the statistical testing of the AUC differences”. What is the test used for the AUC difference? Why is the sample design not accounted for? 6. In the result section on page 11, the authors stated “Russian men were shorter, weighted less and had lower WC compared with Finns”. For this to be valid statistically, a statistical test and p-values should be presented to compare the anthropometric measures between Russian and Finns. Table 1 should be updated with p-values for the comparisons. 7. Similarly, Table 2 and Table 3 should be updated with p-values.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1. As acknowledged by the Authors, the number of participants is quite inadequate; the rate of subjects with >45 years is quite different among groups (61% in Russian, and 31% in Kurdish), and this may influence the results.

Authors’ response: Thank you for your valuable feedback. Our study uses a stratified random sample, therefore the study population is representative of the Russian, Somali, Kurdish and Finnish populations in the six cities where the Maamu Survey has been conducted. Differences in age distributions among the study groups are accounted for in our analyses with age adjustment. In addition to age adjustment, all our analyses have been conducted using inverse probability weights, which have been calculated based on register information from the National Population Register on age, sex, study group, study location and marital status. The purpose of these inverse probability weights is to adjust for non-response bias. This further improves representability of our findings. The population size of the current study is reflected in variance estimates (p-values and confidence intervals). A larger sample size might have resulted in a higher number of statistically significant observations but it does not undermine the validity the findings that were shown to be statistically significant in our study. The statically significant findings in our study would not become insignificant if the sample size would have been larger.

Migrant health examination data with standardized anthropometric measures and blood tests is scarce. To our best knowledge, no previous studies have aimed at determining the best anthropometric measures among Eastern-European, African and Middle-Eastern origin migrants. Additionally, we are unaware of studies using continuous anthropometric measures to examine the association between overweight/obesity indices and type 2 diabetes. The purpose of our study is to raise awareness that commonly used categorical anthropometric measures may misclassify persons with diabetes and that there is a need for large-scale studies on the subject. We discuss the

limitations related to sample size in the limitations section and we are cautious when discussing our results so as not to draw unsupported conclusions.

2. The authors observed that abdominal obesity seems to have less importance in Somali and Kurds, in detecting diabetes, but the contributing role of abdominal obesity in diabetes development is not investigated in this study.

Authors' response: Thank you for raising this very interesting point. As the reviewer kindly pointed out, the contributing role of abdominal obesity in diabetes development is well known among Western populations. We examine the association between categorical overweight and obesity indices, including abdominal obesity and diabetes in Supplementary Table 1 using the following WC categories: normal WC < 94 cm (men)/ < 80 cm (women), overweight 94-102 cm (men)/ 80-88 cm (women), obese > 102 cm (men)/ > 88 cm (women). Our findings show lower accuracy of detection of diabetes among Somali and Kurds for categorical indices compared with continuous anthropometric measures (Table 4). These findings provide support for the critique concerning application of cut-off points for overweight and obesity validated among Caucasians for Middle-Eastern and African origin migrants.

While there is an increasing amount of literature questioning the appropriateness of the currently used obesity cut-off points that have been validated among Caucasian populations for African and Middle-Eastern groups, there are currently no established cut-offs for abdominal obesity among African and Middle-Eastern origin migrants. Furthermore, we are unaware of previous studies that would have used continuous anthropometric measures instead of categorical measures despite the growing awareness that categorical measures are not appropriate for certain migrant groups.

In our paper, the contributing role of abdominal obesity among different study groups can also be to an extent observed in Table 3. From this table, it can be seen that the differences in mean waist circumference among persons without diabetes compared with persons with diabetes are substantially higher among Finns (90.3 cm vs. 104.9 cm) and Russians (87.9 cm vs. 105.8 cm) than among Somali (86.9 cm vs. 92.7 cm) and Kurds (89.9 vs. 96.5 cm). As the reviewer suggested, the association between abdominal obesity and diabetes needs to be further addressed and we will do so in future studies.

3. We also lack information about the duration of diabetes, and the time of diagnosis, i.e. before or after the arrival in a Western country. This information would greatly improve the paper.

Authors' response: Thank you for this valuable suggestion for improving the paper. We have added the following information into the Results section of the paper: "Among persons with type 2 diabetes, mean age upon arrival to Finland was 35 years for Russian, 32 years for Somali and 31 years for Kurds (detailed data not shown). Self-reported mean age at diagnosis was 48 years for Russians, 45 years for Somali, 41 years for Kurds and 45 years for Finns. The difference in age of diagnosis was statistically significant among Kurds (p=0.006) and approached statistical significance for Russians (p=0.058) when comparing with Finns. Type 2 diabetes was diagnosed in Finland in majority of the cases (72% among Russians, 59% among Somali and 70.5% of Kurds)."

We agree that this information is very interesting. The focus of this study is to examine the performance of anthropometric measures that are currently commonly used in simple screening tests for identifying persons at risk for developing future diabetes. It is challenging to explore the contribution of duration of diabetes and time of diagnosis in relation to the association between anthropometrics and onset of diabetes because we do not have data on anthropometrics prior to the

diagnosis of diabetes. The contribution of migration factors are a very interesting matter to be explored in a separate future study.

Reviewer: 2

1. The Maamu survey used samples of 3000 subjects from a stratified random sample. The authors should clarified the stratification procedure.

Authors' response: Thank you for pointing out the need to clarify the stratification procedure. We have added the following sentence into section 2.1 of the Methods section:

"Stratification was based on all combinations of three migrant groups (Russian, Somali and Kurds) and the six cities where the study was conducted (Helsinki, Espoo, Vantaa, Turku, Tampere, Vaasa). Somali migrants were not recruited in the city of Vaasa because at the point of planning the survey, the Somali population size was very low. Stratification was therefore based on altogether 17 combinations (Russian migrants x 6 cities) + (Somali migrants x 5 cities) + (Kurdish migrants x 6 cities). A random sample was drawn in each stratum based on predetermined sample sizes."

2. In the statistical analysis section, the authors stated "all analyses accounted for the stratified sampling and finite population correction". The methods used for achieving this should be specified rather than just specifying the software names.

Authors' response: Thank you for pointing out the need for clarification. We have revised the text the following way:

"All analyses accounted for the stratified sampling and finite population correction and were conducted using the Sudaan 11.0.1 and SAS 9.3 software packages(23). Inverse probability weights, based on register information (age group, sex, study group, study location, and marital status), were used to correct for the effects of non-response and different sampling probabilities in all of the analyses(24). The regression analyses were based on the generalized estimating equations, and all variance estimates on linearization(25)."

3. On page 9, last line, the authors mentioned linear logistic regression was used to calculate age-adjusted mean values and 95% CI for continuous variables. This does not make sense.

Authors' response: Thank you for pointing out this typing error. The intention was to write about using linear regression for continuous measures. We have corrected this error the text.

4. On page 10, first two lines, the authors mentions regression analyses by study group stratified by sex. This is very confusing and needs to be clarified. In the result section, I did not see any results of the AUC by sex.

Authors' response: Thank you for pointing out the need for revision. We have revised the text the following way:

"All regression analyses were stratified by study group and sex, with exception for estimation of age-adjusted mean anthropometric measures according to the presence of type 2 diabetes. All regression analyses were conducted using predictive margins based on regression models(26)… ROC analyses were performed for each study group separately."

5. The last sentence of the method section, “With exception of inverse probability weights, the sampling design was not accounted for in the statistical testing of the AUC differences”. What is the test used for the AUC difference? Why is the sample design not accounted for?

Authors’ response: Thank you for pointing out the need for this clarification. As mentioned in the text, inverse probability weights, which reduce non-response bias were included in all of the analyses. The complex stratified sampling design was based on 17 combinations in the Maamu Survey as mentioned in our response to your first comment. This stratified sampling design accounts for different sampling probabilities.

Different sampling probabilities and non-response bias (inverse probability weights) were accounted for when calculating the AUC but it was not possible to account for sampling design when calculating confidence intervals for AUC and for calculating the p-values for the differences in AUC values for BMI, WC and WHR in comparison with WHtR. The reason for not accounting for sampling design in these analyses is that we are unaware of any accessible statistical tools for assessing the differences of AUC in complex survey data. Not accounting for study design may have led to narrower confidence intervals and slightly lower p-values. In our study, we report no statistically significant differences in the performance of anthropometric measures with exception for BMI among Finns ($p<0.001$). Despite the possibility that the p-value would slightly increase if we would have been able to take the sampling design into account, the finding would have nonetheless remained statistically significant.

We clarified this in the text in the following way:

“Stratified sampling design based on the 17 combinations of migrant groups and study locations were accounted for when calculating the AUC values but not when calculating confidence intervals for AUC and p-values for the difference in the performance of anthropometric measures within each migrant group. The reason for not accounting for stratified sampling design in these analyses is that we are unaware of any accessible statistical tools for assessing the differences of AUC in complex survey data.”

6. In the result section on page 11, the authors stated “Russian men were shorter, weighted less and had lower WC compared with Finns”. For this to be valid statistically, a statistical test and p-values should be presented to compare the anthropometric measures between Russian and Finns. Table 1 should be updated with p-values for the comparisons.

Authors’ response: Thank you for the comment. We added * symbols for p-values to indicate significance. We chose *-symbols so that the table is easier to read.

7. Similarly, Table 2 and Table 3 should be updated with p-values.

Authors’ response: The tables were revised according to reviewer’s suggestion.

VERSION 2 – REVIEW

REVIEWER	Hongyan Xu Augusta University, United States
REVIEW RETURNED	05-Mar-2018

GENERAL COMMENTS	This is a revision. The authors have addressed my previous concerns sufficiently.

REVIEWER	Saula de Kreutzenberg University of Padova, Italy
REVIEW RETURNED	08-Mar-2018

GENERAL COMMENTS	Authors have considered the reviewer's suggestions and integrated the paper in accordance.